

09/657,289 # 12

(19) RÉPUBLIQUE FRANÇAISE
 INSTITUT NATIONAL
 DE LA PROPRIÉTÉ INDUSTRIELLE
 PARIS

(11) N° de publication : 2 693 475

(à n'utiliser que pour les commandes de reproduction)

(21) N° d'enregistrement national : 92 08427

(51) Int Cl⁶ : C 12 Q 1/68, C 12 N 15/63, 15/81, C 07 H 21/04

(12)

DEMANDE DE BREVET D'INVENTION

A1

(22) Date de dépôt : 08.07.92.

(71) Demandeur(s) : RHONE-POULENC RORER (S.A.) — FR.

(30) Priorité :

(72) Inventeur(s) : Menart Sandrine et Bolotin Monique.

(43) Date de la mise à disposition du public de la demande : 14.01.94 Bulletin 94/02.

(73) Titulaire(s) :

(56) Liste des documents cités dans le rapport de recherche préliminaire : Se reporter à la fin du présent fascicule.

(74) Mandataire : Rhône-Poulenc Rorer S.A.

(54) Procédé d'identification et/ou de clonage de promoteurs transcriptionnels, et utilisation de ces promoteurs pour l'expression de gènes.

(57) L'invention concerne un procédé d'identification et/ou de clonage de promoteurs transcriptionnels fonctionnels dans un hôte donné, les promoteurs ainsi obtenus et leur utilisation pour l'expression de gènes.

FR 2 693 475 - A1



PROCEDE D'IDENTIFICATION ET/OU DE CLONAGE DE PROMOTEURS
TRANSCRIPTIONNELS, ET UTILISATION DE CES PROMOTEURS POUR
L'EXPRESSION DE GENES

La présente invention concerne le domaine du génie génétique. Plus 5 particulièrement, l'invention concerne un procédé pour l'identification et le clonage de promoteurs transcriptionnels.

Les progrès accomplis dans le domaine de la biologie moléculaire ont permis de modifier des microorganismes pour leur faire produire des protéines hétérologues. En particulier, de nombreuses études génétiques ont 10 porté sur la bactérie *E. coli* et sur la levure *Saccharomyces*. Toutefois, l'application industrielle de ces nouveaux modes de production est encore limitée, en particulier par les problèmes d'efficacité d'expression des gènes 15 dans ces microorganismes recombinés. Aussi, dans le but d'augmenter les performances de ces systèmes de production, il est nécessaire de pouvoir disposer de promoteurs forts, permettant d'obtenir des niveaux élevés 20 d'expression de protéines hétérologues. Il peut également être souhaitable de disposer d'un grand nombre de promoteurs variés, actifs dans différentes conditions physiologiques.

Chez *E.coli*, certains promoteurs forts ont été décrits (promoteurs des 25 opérons tryptophane et lactose). De même, chez la levure *S. cerevisiae*, des études ont porté sur des promoteurs dérivés de gènes impliqués dans la glycolyse. On peut citer notamment les travaux sur le promoteur du gène de la 3-phosphoglycéate kinase PGK (Dobson et al., Nucleic Acid Res. 10, 1982, 2625; Hitzeman et al., Nucleic Acid Research 1982, 7791), sur celui du gène 30 de la glyceraldéhyde-3-phosphate deshydrogénase GAPDH (Holland et al., J.Biol.Chem. 254, 1979, 9839; Musti et al., Gene 25, 1983, 133), sur celui du gène de l'alcool deshydrogénase 1 ADH1 (Bennentzen et al., J.Biol.Chem. 257, 1982, 3018; Denis et al., J.Biol.Chem. 25, 1983, 1165), sur celui du gène de l'enolase 1 ENO1 (Uemura et al., Gene 45, 1986, 65), sur celui du gène GAL1/GAL10 (Johnston et Davis, Mol. Cell. Biol. 4 (1984) 1440) ou sur celui du gène CYC1 (Guarente et Ptashne, PNAS 78 (1981) 2199).

Toutefois, il est clair que le génome de ces cellules n'a pu être étudié complètement, et qu'il renferme potentiellement d'autres promoteurs forts non encore identifiés. De même, de plus en plus, des procédés recombinants

2693475

2

visent aujourd'hui à utiliser de nouveaux types de cellules hôtes (cellules animales, végétales, levures, champignons ou bactéries).

Ainsi, des outils génétiques ont été développés afin de se servir de la levure *Kluyveromyces* comme cellule hôte pour la production de protéines recombinantes (EP 361 991). Des études ont également été entreprises avec *Bacillus subtilis* (Saunders et al., J. Bacteriol. 169 (1987) 2917), avec levure de brasserie (EP 201 239), ou encore avec les levures *Pichia pastoris* (EP 344 459), *Schizosaccharomyces pombe* (EP 385 391) ou *Schwanniomyces* (EP 394 538).

10 Cependant, les promoteurs utilisés dans ces nouveaux systèmes ne sont pas encore optimisés. En particulier, il s'agit souvent de promoteurs hétérologues, c'est-à-dire provenant d'autres microorganismes, ce qui peut engendrer différents inconvénients, et notamment limiter l'activité du promoteur à cause de l'absence de certains éléments de la machinerie transcriptionnelle (par exemple de trans-activateurs), présenter une certaine toxicité pour la cellule hôte due à une absence de régulation, ou affecter la stabilité du vecteur.

15 20 Dans ces conditions, l'exploitation de ces systèmes de production impose de posséder des promoteurs adaptés, préférentiellement homologues, régulables et forts. Il est donc important de pouvoir disposer d'un moyen efficace, simple et fiable permettant de rechercher et d'isoler des promoteurs à partir de tout hôte cellulaire.

25 La demanderesse a maintenant mis au point, et ceci constitue un objet de la présente invention, un procédé particulièrement avantageux permettant d'identifier et de cloner des promoteurs transcriptionnels. Plus précisément, l'invention concerne un procédé d'identification et/ou de clonage de promoteurs transcriptionnels fonctionnels dans un hôte donné, selon lequel on effectue les étapes suivantes :

- on prépare une banque d'ADN d'une cellule de départ déterminée,
- 30 - on réalise une mutagénèse insertionnelle par transposition sur cette banque permettant d'insérer dans l'ADN de la cellule de départ un transposon portant un gène reporteur dont l'expression dans l'hôte donné peut être mise en évidence,
- on sélectionne le(s) clone(s) exprimant, dans des conditions de culture déterminées, ledit gène reporteur dans l'hôte donné, et,

2693475

3

- on isole le fragment d'ADN comprenant la région promotrice du (des) clone(s) ainsi sélectionné(s).

Le procédé de clonage selon l'invention est particulièrement avantageux puisqu'il peut être appliqué à toute cellule de départ, et qu'il n'est pas nécessaire de connaître la fonction et/ou la structure du gène naturellement exprimé par ce promoteur. Ce procédé permet ainsi de cloner des promoteurs parfaitement inconnus. Par ailleurs, un autre avantage du procédé de l'invention est qu'il permet de cloner des promoteurs actifs dans des conditions de culture déterminées, et de ce fait d'orienter la sélection des promoteurs en fonction par exemple d'un type de régulation souhaité. Le procédé de l'invention est également avantageux puisqu'il peut permettre l'identification rapide du gène placé sous le contrôle du promoteur isolé par détermination d'un court fragment de séquence au niveau de la jonction entre le transposon et le fragment d'ADN de la souche de départ au moyen d'un oligonucléotide spécifique du transposon.

En ce qui concerne la première étape du procédé de l'invention, elle est applicable à tout type de cellule, et notamment aux cellules eucaryotes et procaryotes.

Parmi les cellules eucaryotes, on peut citer les cellules végétales, animales, les levures, ou les champignons. En particulier, s'agissant de levures, on peut citer les levures du genre *Saccharomyces*, *Kluyveromyces*, *Pichia*, *Schwanniomyces*, ou *Hansenula*. S'agissant de cellules animales, on peut citer les cellules COS, CHO, Cl27, etc. Parmi les champignons susceptibles d'être utilisés dans la présente invention, on peut citer plus particulièrement *Aspergillus* ssp. ou *Trichoderma* ssp. Comme hôtes procaryotes, on peut utiliser les bactéries telles que *Escherichia coli*, ou celles appartenant aux genres *Corynebacterium*, *Bacillus*, ou *Streptomyces*.

Le choix de la cellule de départ peut être dicté par différents critères (facilité de manipulation, fort niveau d'expression de protéines, analyse d'une cellule peu étudiée, etc).

La banque d'ADN de la cellule de départ peut être une banque génomique ou une banque d'ADNc. Celles-ci peuvent être préparées par toute technique connue de l'homme du métier. Notamment, s'agissant d'une banque génomique, elle peut être préparée par rupture des cellules, digestion

2693475

4

de l'ADN au moyen d'enzymes de restriction et insertion de l'ADN ainsi digéré dans un vecteur. Généralement, on utilise des conditions de digestion (enzymes, temps de réaction, etc) permettant d'obtenir des fragments d'une taille d'environ 4 à 8 kb. S'agissant d'une banque d'ADNc, elle peut être préparée à partir des ARN, par transcription inverse.

Pour des raisons de commodité, la banque est généralement préparée chez *E.coli*, dans des vecteurs de type pBR322. Par ailleurs, pour permettre le transfert de la banque depuis la souche initiale (*E.coli*) vers l'hôte dans lequel l'expression est recherchée, il est possible d'utiliser soit des vecteurs navettes capables de réPLICATION dans les 2 types d'hôtes, soit de construire des transposons possédant les éléments nécessaires à la réPLICATION et la sélection dans l'hôte dans lequel l'expression est recherchée. Cette deuxième alternative présente l'avantage de ne pas utiliser des vecteurs de taille trop grande pour réaliser la banque.

En ce qui concerne la deuxième étape du procédé de l'invention, le gène reporteur est un gène dont l'expression peut être mise en évidence dans l'hôte donné. Il peut s'agir d'un gène induisant une coloration des souches (gène lacZ notamment) ou leur conférant un caractère luminescent (gène luxA notamment), ou encore d'un gène conférant une résistance à la souche [gène conférant la résistance au G418 (gène aph) ou au cuivre (gène CUP1) notamment]. Il peut également s'agir d'un gène indispensable à la croissance de la souche dans certaines conditions de culture, etc.

A titre d'exemple de gène reporteur particulièrement avantageux pour les levures *Klactis lac4*, le gène lacZ d'*E.coli* codant pour β -galactosidase permet une sélection sur milieu lactose (Chen et al., J. Basic Microbiol. 28 (1988) 211).

Le gène reporteur doit être présent dans le transposon dans une forme ne permettant son expression que par fusion traductionnelle avec un gène de l'hôte donné. Notamment, il ne doit pas être précédé de codon stop ou d'autres signaux qui pourraient empêcher son expression.

A titre d'exemples de transposons utilisables dans le cadre de l'invention, on peut citer plus particulièrement le transposon MiniMu et les transposons Tn (Tn3, Tn5, Tn7, Tn10, Tnbla, etc).

2693475

5

La troisième étape du procédé de l'invention consiste à sélectionner le(s) clone(s) exprimant le gène reporteur dans l'hôte donné, dans des conditions de culture déterminées. Dans le cas où la banque est réalisée dans une cellule différente de l'hôte donné dans lequel l'expression est recherchée,

5 la banque doit dans un premier temps être transférée dans l'hôte donné, avant de procéder à la sélection des clones exprimant le gène reporteur. Dans ce cas, une première sélection est effectuée pour isoler les transformants ayant reçu l'ADN de la banque. Cette sélection peut être effectuée en utilisant tout marqueur d'auxotrophie ou de résistance. Le transfert de l'ADN de la

10 banque peut être réalisé par toute technique connue de l'homme du métier, et notamment par transformation, électroporation, ou conjugaison. S'agissant de la transformation chez la levure, différents protocoles ont été décrits dans l'art antérieur. En particulier, elle peut être réalisée en traitant les cellules entières en présence d'acétate de lithium et de polyéthylène glycol selon la

15 technique décrite par Ito et al. (*J. Bacteriol.* 153 (1983) 163-168), ou en présence d'éthylène glycol et de diméthylsulfoxyde selon la technique de Durrens et al. (*Curr. Genet.* 18 (1990) 7). Un protocole alternatif a également été décrit dans la demande de brevet EP 361 991. Il est aussi possible de transformer les levures par électroporation, selon la méthode décrite par

20 Karube et al. (*FEBS Letters* 182 (1985) 90). La sélection des clones exprimant le gène reporteur dans l'hôte donné est réalisée après transformation, dans différentes conditions de culture selon les caractéristiques du promoteur recherché (promoteur régulable, promoteur fort, etc). Par ailleurs, selon le gène reporteur utilisé, les clones possédant une région promotrice active dans

25 les conditions de la sélection sont ceux générant une certaine coloration ou une luminescence, ceux présentant une résistance à un composé donné (gène de résistance) ou encore ceux capables de pousser dans certains milieux (gènes complémentant une auxotrophie tels que les gènes lacZ, URA3 ou LEU2, notamment), etc.

30 Dans un mode avantageux de l'invention, il est en outre possible d'effectuer simultanément la sélection des transformants et la sélection des clones exprimant le gène reporteur.

Lors de la quatrième étape du procédé de l'invention, l'isolement du fragment d'ADN comprenant la région promotrice est généralement réalisée

2693475

6

par excision au moyen d'enzymes de restriction, suivie éventuellement d'une digestion par des nucléases ou d'une synthèse chimique pour reconstituer une partie manquante. Préférentiellement, l'excision est précédée d'une étape de délimitation de la région promotrice portée par le fragment d'ADN. Celle-ci est généralement réalisée en plusieurs temps :

5 - détermination d'une carte de restriction de l'insert porté par le clone. Cette étape est néanmoins facultative,

10 - localisation de l'extrémité 3' de la région promotrice : Cette étape consiste à identifier le codon ATG. Elle est généralement réalisée par séquençage de la jonction entre le transposon et le fragment d'ADN de la souche de départ. Ce bout de séquence peut ensuite permettre, par recherche d'homologies avec des séquences connues, de localiser l'ATG. Si la recherche d'homologie ne donne pas de résultat, l'ATG peut être identifié par séquençage complet du fragment d'ADN de la souche de départ portant la 15 région promotrice.

15 - localisation de l'extrémité 5' de la région promotrice : Celle-ci peut être mise en évidence par des délétions séquentielles sur le fragment d'ADN de la souche de départ portant la région promotrice et sous-clonage dans des vecteurs d'expression.

20 Les promoteurs obtenus par le procédé de l'invention peuvent être utilisés pour l'expression de tout gène soit dans la cellule de départ, soit dans l'hôte donné, soit encore dans tout autre cellule dans lesquels ils sont fonctionnels. Ils peuvent notamment être utilisés pour l'expression de gènes codant pour des protéines recombinantes telles que par exemple des 25 protéines d'intérêt pharmaceutique ou agroalimentaire, ou également pour l'expression de gènes de biosynthèse de métabolites, tels que par exemple des gènes de biosynthèse d'acides aminés, de vitamines, d'antibiotiques, etc.

30 Par ailleurs, ces fragments d'ADN peuvent également être modifiés préalablement à leur utilisation, par exemple pour préparer des promoteurs portables, pour les adapter à l'expression dans un type particulier de vecteur ou d'hôte, pour réduire leur taille, pour augmenter leur activité de promoteur de la transcription, pour générer des promoteurs inducibles, améliorer leur niveau de régulation, ou encore changer la nature de leur régulation. De telles modifications peuvent être effectuées par exemple par mutagénèse *in*

2693475

7

vitro, par introduction d'éléments additionnels de contrôle ou de séquences synthétiques, ou par des délétions ou des substitutions des éléments originels de contrôle.

L'invention a également pour objet un procédé de production d'une protéine recombinante par expression d'une séquence d'ADN codant pour ladite protéine dans un hôte cellulaire, selon lequel l'expression de la séquence d'ADN est sous le contrôle d'un promoteur obtenu par le procédé décrit ci-dessus.

D'autres avantages de la présente invention apparaîtront à la lecture des exemples suivants, qui doivent être considérés comme illustratifs et non limitatifs.

LEGENDE DES FIGURES

Figure 1 : Préparation du transposon Mini Mu MudIZK1.

Figure 2 : Carte de restriction du transposon Mini Mu MudIZK1.

Figure 3 : Séquence de la jonction entre le promoteur du gène PGK de *K.lactis* et le gène lacZ de *E.coli* (mini mu MudIZK1) sur le clone 4B3.

Figure 4 : Séquence de la jonction entre le promoteur du gène ADHII de *K.lactis* et le gène lacZ de *E.coli* (mini mu MudIZK1) sur le clone 10B1.

Figure 5 : Carte de restriction de l'insert du clone 1D12.

Figure 6 : Séquence de la jonction entre le promoteur du gène PDC1 de *K.lactis* et le gène lacZ de *E.coli* (mini mu MudIZK1) sur le clone 1D12.

TECHNIQUES GENERALES DE CLONAGE.

Les méthodes classiquement utilisées en biologie moléculaire telles que les extractions préparatives d'ADN plasmidique, la centrifugation d'ADN plasmidique en gradient de chlorure de césum, l'électrophorèse sur gels d'agarose ou d'acrylamide, la purification de fragments d'ADN par électroélution, les extraction de protéines au phénol ou au phénol-chloroforme, la précipitation d'ADN en milieu salin par de l'éthanol ou de l'isopropanol, la transformation dans *Escherichia coli* etc, sont bien connues de l'homme de métier et sont abondamment décrites dans la littérature [Maniatis T. et al., "Molecular Cloning, a Laboratory Manual", Cold Spring Harbor

2693475

8

Laboratory, Cold Spring Harbor, N.Y., 1982; Ausubel F.M. et al. (eds), "Current Protocols in Molecular Biology", John Wiley & Sons, New York, 1987].

5 Les enzymes de restriction ont été fournies par New England Biolabs (Biolabs) ou Pharmacia et sont utilisées selon les recommandations des fournisseurs.

Les plasmides de type pBR322, pUC et les phages de la série M13 sont d'origine commerciale (Bethesda Research Laboratories).

10 Pour les ligatures, les fragments d'ADN sont séparés selon leur taille par électrophorèse en gels d'agarose ou d'acrylamide, extraits au phénol ou par un mélange phénol/chloroforme, précipités à l'éthanol puis incubés en présence de l'ADN ligase du phage T4 (Boehringer) selon les recommandations du fournisseur.

15 Le remplissage des extrémités 5' proéminentes est effectué par le fragment de Klenow de l'ADN Polymérase I d'*E. coli* (Boehringer) selon les spécifications du fournisseur. La destruction des extrémités 3' proéminentes est effectuée en présence de l'ADN Polymérase du phage T4 (Biolabs) utilisée selon les recommandations du fabricant. La destruction des extrémités 5' proéminentes est effectuée par un traitement ménagé par la nucléase S1.

20 La mutagénèse dirigée *in vitro* par oligodéoxynucléotides synthétiques est effectuée selon la méthode développée par Taylor et al. [Nucleic Acids Res. 13 (1985) 8749-8764].

25 L'amplification enzymatique de fragments d'ADN par la technique dite de PCR [Polymérase-catalyzed Chain Reaction, Saiki R.K. et al., Science 230 (1985) 1350-1354; Mullis K.B. et Falona F.A., Meth. Enzym. 155 (1987) 335-350] est effectuée en utilisant un "DNA thermal cycler" (Perkin Elmer Cetus) selon les spécifications du fabricant.

30 La vérification des séquences nucléotidiques est effectuée par la méthode développée par Sanger et al. [Proc. Natl. Acad. Sci. USA, 74 (1977) 5463-5467].

Les transformations de *K. lactis* sont effectuées par toute technique connue de l'homme de l'art, et dont un exemple est donné dans le texte.

35 Sauf indication contraire, les souches bactériennes utilisées sont *E.coli* DH1 (Hanahan D., J. Mol. Biol. /166 (1983) 557) ou *E.coli* JM109::(Mucts) (Daignan-formier et Bolotin-Fukuhara, Gene 62 (1988) 45).

2693475

9

Les souches de levures utilisées appartiennent aux levures bourgeonnantes et plus particulièrement aux levures du genre *Kluyveromyces*. Les souches *K. lactis* 2359/152 et *K. lactis* SD6 ont été particulièrement utilisées.

5 Les souches de levures transformées par les plasmides sont cultivées en erlenmeyers ou en fermentateurs pilotes de 21 (SETRIC, France) à 28°C en milieu riche (YPD: 1% yeast extract, 2% Bactopeptone, 2% glucose; ou YPL: 1% yeast extract, 2% Bactopeptone, 2% lactose) sous agitation constante.

EXEMPLES

10 EXEMPLE DE CLONAGE DE PROMOTEURS DE *KLUYVEROMYCES* EN UTILISANT COMME GENE REPORTEUR LE GENE LACZ DE *E.COLI* ET COMME SYSTEME DE MUTAGENESE LA TRANSPOSITION.

1. Préparation du transposon Mini Mu MudIZK1 (figures 1 et 2)

15 Le Mini Mu MudIZK1 a été construit à partir du Mini Mu MudIZZ1 décrit par Daignan-Fornier et Bolotin-Fukuhara (Gene 62 (1988) 45). Il a été obtenu en substituant l'origine de réPLICATION du mini transposon MudIZZ1 par une origine de réPLICATION fonctionnelle dans *Kluyveromyces* : l'origine de réPLICATION du plasmide pKD1 (EP231 435).

1.1. Construction d'une cassette portant l'origine de réPLICATION du plasmide pKD1 (fragment S11).

20 Afin de faciliter les manipulations ultérieures, le fragment S11 (portant l'origine de réPLICATION du plasmide pKD1) a été mis sous forme d'une cassette NotI. Pour cela, un dérivé du plasmide pUC18 a été construit, dans lequel les sites externes du multisite de clonage (sites HindIII et EcoRI) ont été changés en sites NotI. Cela a été fait par digestion avec l'enzyme 25 correspondante, action de l'enzyme de Klenow et ligature avec un oligonucléotide synthétique correspondant à un site NotI [oligo d(AGCGGCCGCT); Biolabs]. Le plasmide obtenu est désigné pGM67. Le fragment S11 de 960 pb obtenu par digestion par l'enzyme Sau3A du plasmide KEp6 (Chen et al, Nucl. Acids Res. 14 (1986) 4471) a ensuite été 30 inséré au site compatible BamHI du plasmide pGM67. Le plasmide ainsi

2693475

10

obtenu, désigné pGM68, contient sous forme d'une cassette NotI le fragment S11.

1.2. Suppression de l'origine de réPLICATION 2 μ du transposon MudIZZ1.

5 Le plasmide pGM15 portant le mini Mu MudIZZ1 (Daignan-Fornier et Bolotin-Fukuhara précitée) a été déléte des régions 2 μ par digestion au moyen de l'enzyme SalI. Le site SalI unique ainsi obtenu a ensuite été transformé en site NotI par ligature d'un oligonucléotide synthétique correspondant à un site NotI après action de l'enzyme de Klenow. Le 10 plasmide résultant est appelé pGM59.

1.3. Insertion du fragment S11

La cassette NotI portant l'origine de réPLICATION du plasmide pKD1 (fragment S11), provenant du plasmide pUC18 modifié a ensuite été introduite au site NotI unique du plasmide pGM59.

15 Le plasmide obtenu, désigné pGM83, porte un mini Mu, appelé MudIZK1, qui est adapté à la levure *Kluyveromyces lactis*, ainsi qu'une copie fonctionnelle du gène LEU2 de *S.cerevisiae* capable de complémenter une mutation leu2 chez *K.lactis* (Kämper et al, Curr. Genet. 19 (1991) 109) (figure 2).

20 **2. Introduction du Mini Mu MudIZK1 dans la souche *E.coli* portant le Mu helper JM109::(Mucts) : Obtention de la souche JM109::(Mucts)::(MudIZK1).**

25 La souche JM109::(Mucts) a été transformée par le plasmide pGM83 contenant le mini mu MudIZK1 en présence de chlorure de calcium. Après transformation, la transposition a été induite par choc thermique selon la technique décrite par Castilho et al. (J. Bacteriol. 158 (1984) 488). Le lysat phagique obtenu après induction est ensuite utilisé pour surinfecter la souche JM109::(Mucts). La souche JM109::(Mucts) étant recA, l'ADN linéaire encapsidé par le phage ne peut se refermer pour donner un plasmide 30 répliquatif. Les intégrants [souche JM109::(Mucts)::(MudIZK1)] sont donc

2693475

11

sélectionnés comme clones chloramphénicol résistants (Cm^R), ampicilline sensibles (Amp^S).

3. Préparation de la banque génomique de *K.lactis* dans *E.coli* DH1

L'ADN de haut poids moléculaire a été préparé à partir de la souche 5 *K.lactis* 2359/152, et digéré partiellement par l'enzyme Sau3A. Les fragments d'une taille de 4 à 8 kb ont été récupérés sur gel d'agarose LMP ("Low Melting Point", SEAKEM) et clonés dans le plasmide pBR322 linéarisé par BamHI et déphosphorylé par action de la phosphatase intestinale de veau (Biolabs). 35 pools de 1000 colonies ampicilline résistantes (Amp^R) et 10 tétracycline sensibles (Tet^S) dans *E.coli* DH1 ont ainsi été réalisés.

4. Préparation de la banque de fusion

4.1. Introduction de la banque génomique de *K.lactis* dans la souche JM109::(Mucts)::(MudIIZK1).

L'ADN plasmidique de chaque pool réalisé dans DH1 est extrait 15 (Maniatis). Cet ADN est ensuite utilisé pour transformer la souche JM109::(Mucts)::(MudIIZK1) en présence de chlorure de calcium. Pour être représentatif des 1000 colonies contenues dans chaque pool de la banque génomique, plus de 3000 clones par pool ont été récupérés dans la souche JM109::(Mucts)::(MudIIZK1) permettant la transduction.

20 4.2. Transposition du Mini Mu MudIIZK1

La banque de fusion est réalisée par transposition extensive du Mini Mu MudIIZK1 sur les plasmides formant la banque d'ADN génomique de *K.lactis*. Les mini-muductions ont été faites selon le protocole décrit par Castilho et al. (J. Bacteriol. 158 (1984) 488) et les transductants ont été 25 sélectionnés sur milieu sélectif LBAC (milieu LB (Gibco, BRL) supplémenté avec 50 mg/l d'ampicilline et 30 mg/l de chloramphénicol), le marqueur Amp^R étant apporté par le plasmide, et le marqueur Cm^R par le mini-mu. Pour chaque pool, des transpositions sont faites en série, et entre 10 000 et 20 000 transductants sont récupérés par pool. L'ADN des transductants est 30 ensuite extrait d'une préparation de 100 ml, purifié par précipitation au polyéthylène glycol (Maniatis et al, 1989) et resuspendu dans 100µl d'eau. Cet

2693475

12

ADN a ensuite été utilisé pour transformer *K.lactis* et pour cloner des promoteurs.

5. Recherche de promoteurs chez *K.lactis*

L'ADN de fusion préparé ci-dessus a été utilisé pour transformer, par électroporation, une souche réceptrice de *K.lactis*. Cette souche réceptrice, désignée SD6, porte les mutations leu2 (correspondant au marqueur de sélection du mini-mu MudIZK1) et lac4-8. Cette dernière mutation empêche la souche de pousser sur un milieu contenant du lactose comme seule source de carbone, mais elle peut être complémentée par la surexpression du gène lacZ d'*E.coli* codant pour la β -galactosidase (Chen et al., J. Basic Microbiol. 28 (1988) 211). De ce fait, l'expression d'une protéine fusionnée à la β -galactosidase doit permettre la croissance de la souche SD6 sur lactose après transformation. Ce cible positif a été utilisé pour sélectionner rapidement des clones portant des promoteurs forts.

15 5.1. Construction de la souche réceptrice *K.lactis* SD6.

Cette souche doit porter les marqueurs suivants :

- marqueur d'auxotrophie complémenté par le marqueur du mini mu (LEU2)
- marqueur Ade⁺ pour éviter des problèmes de lecture sur milieu Xgal dus à la coloration rouge de l'allèle Ade⁻,
- marqueur lac4 (muté pour la β -galactosidase). L'utilisation de ce marqueur doit permettre, dans le cas d'un promoteur fort et donc d'une protéine de fusion fortement exprimée, de complémer le phénotype lac de la souche.

25 La souche SD6 (Chen et al., Mol. Gen. Genet. 233 (1992) 97) a été obtenue par croisement de la souche CXJ1-7A (a, lac4-8, uraA, ade1-1, K1, K2, pKD1) (Chen et Fukuhara, Gene 69 (1988) 18187) avec la souche AWJ-137 (leu2, trp1, homothallique) (Kämper et al, Curr. Genet. 19 (1991) 109), et sélection des spores ayant le génotype ADE⁺, uraA, leu2, lac4-8. Comme les 30 spores obtenues n'étaient pas capables de régénérer après transformation par protoplastes, un croisement retour a été fait avec la souche CXJ1-7A. Après sporulation en masse, les spores du génotype choisi ont été testées par

2693475

13

transformation au chlorure de lithium avec le plasmide KEp6 selon une technique dérivée de celle décrite par Ito et al. (J. Bacteriol. 153 (1983) 163) (la concentration en LiCl est de 20 mM, soit 10 fois moins que celle utilisée par Ito pour *S.cerevisiae*). La souche CXJ1-7A a servi de témoin de transformation.

5 La souche SD6, sélectionnée sur ces critères, se transforme correctement : 1 à 3. 10⁴ transformants par µg d'ADN; et les transformants ont une stabilité satisfaisante : 30 à 40% des colonies gardent le phénotype [Ura⁺] après 6 générations en milieu non sélectif.

5.2 Isolement de promoteurs forts

10 La souche SD6 a été transformée par électroporation selon Becker et Guarante (in Methods in Enzymology vol194 (1991) 182) (appareil Jouan; 2500 V/cm; 80-100 ng d'ADN/Transformation) avec l'ADN de 11 pools de transductants obtenus en 4. (correspondant à une banque de 11000 clones dans *E.coli*). Après 5 heures de régénération en milieu YPD (extrait de levure 15 : 10g/l; peptone : 10g/l; glucose 20g/l), les cellules ont été étalées sur milieu minimum lactose. Les transformants capables de pousser sur lactose ont été restriés et, pour chaque clone, le plasmide a été extrait, amplifié dans *E.coli*, et, après vérification rapide de la carte de restriction du vecteur et du mini-mu, utilisé pour retransformer la levure SD6. Après retransformation, 20 différents clones de *K.lactis* capables de pousser sur lactose ont été mis en évidence. Chacun de ces clones porte une région promotrice capable d'exprimer un gène hétérologue dans la levure *K.lactis*. Parmi les clones ainsi identifiés, plusieurs ont été analysés par restriction et par séquençage des jonctions entre la protéine de *K.lactis* et la β-galactosidase. En effet, la 25 séquence des jonctions à partir de l'extrémité lacZ du mini-mu (séquence double brin) peut être déterminée par séquençage au moyen d'un oligonucléotide complémentaire des extrémités du mini-mu. Il est notamment possible d'utiliser les oligonucléotides ayant la séquence suivante :

- 30 (i) oligonucléotide amont : 5'-CTGTTTCATTGAAGCGCG-3'
(ii) oligonucléotide aval : 5'-CCCACCAATCTAATCCC-3'

2693475

14

Dans cet exemple, la séquence a été réalisée au moyen de l'oligonucléotide (i) situé à -59 nucléotides de la jonction. Cette étape de séquençage permet de vérifier l'insertion du mini-mu en phase dans une phase de lecture ouverte (on attend une ORF dans le cadre -1) et, lorsque c'est 5 le cas, d'analyser la séquence protéique déduite de la séquence nucléotidique par comparaison avec les séquences de banques de protéines d'autres levures ou eucaryotes (Genbank, MIPS, EMBL, etc).

L'analyse de la séquence des jonctions a permis de montrer que le 10 clone 4B3 portait le promoteur du gène de la phospho-glycérate kinase de *K.lactis* (figure 3), que le clone 10B1 portait le promoteur du gène de l'alcool déshydrogénase II de *K.lactis* (figure 4), et que le clone 1D12 portait le promoteur du gène de la pyruvate décarboxylase de *K.lactis* (figures 5 et 6).

2693475

15

REVENDICATIONS

1. Procédé d'identification et/ou de clonage de promoteurs transcriptionnels fonctionnels dans un hôte donné, caractérisé en ce que l'on effectue les étapes suivantes :

- 5 - on prépare une banque d'ADN d'une cellule de départ déterminée,
 - on réalise une mutagénèse insertionnelle par transposition sur cette banque permettant d'insérer dans l'ADN de la cellule de départ un transposon portant un gène reporteur dont l'expression dans l'hôte donné peut être mise en évidence,
- 10 - on sélectionne le(s) clone(s) exprimant, dans des conditions de culture déterminées, ledit gène reporteur dans l'hôte donné, et,
 - on isole le fragment d'ADN comprenant la région promotrice du (des) clone(s) ainsi sélectionné(s).

15 2. Procédé selon la revendication 1 caractérisé en ce que la cellule de départ est une cellule eucaryote ou procaryote.

15 3. Procédé selon la revendication 1 caractérisé en ce que la banque de la cellule de départ est une banque génomique préparée par rupture des cellules, digestion de l'ADN au moyen d'enzymes de restriction et insertion de l'ADN ainsi digéré dans un vecteur.

20 4. Procédé selon la revendication 3 caractérisé en ce que l'on utilise des conditions de digestion permettant d'obtenir des fragments d'une taille d'environ 4 à 8 kb.

25 5. Procédé selon la revendication 1 caractérisé en ce que la banque de la cellule de départ est une banque d'ADNc préparée à partir des ARN par transcription inverse.

30 6. Procédé selon la revendication 1 caractérisé en ce que le gène reporteur est un gène dont l'expression induit une coloration de l'hôte donné, ou lui confère un caractère luminescent, ou lui confère une résistance, ou est un gène indispensable à la croissance de l'hôte donné dans certaines conditions de culture.

2693475

16

7. Procédé selon la revendication 6 caractérisé en ce que le gène reporteur est choisi parmi le gène LacZ, LuxA, CUP1 et aph.

8. Procédé selon la revendication 6 ou 7 caractérisé en ce que le gène reporteur est présent dans le transposon dans une forme ne permettant son expression que par fusion traductionnelle avec un gène de l'hôte donné.

5 9. Procédé selon la revendication 1 caractérisé en ce que le transposon est choisi parmi le transposon mini mu et les transposons Tn..

10. Utilisation d'un promoteur obtenu par le procédé selon l'une quelconque des revendications 1 à 9 pour l'expression de gènes.

10 11. Utilisation selon la revendication 10 pour l'expression de gènes codant pour des protéines recombinantes telles que par exemple des protéines d'intérêt pharmaceutique ou agroalimentaire, ou également pour l'expression de gènes de biosynthèse de métabolites, tels que par exemple des gènes de biosynthèse d'acides aminés, de vitamines, ou d'antibiotiques.

15 12. Procédé de production d'une protéine recombinante par expression d'une séquence d'ADN codant pour ladite protéine dans un hôte cellulaire caractérisé en ce que l'expression de la séquence d'ADN est sous le contrôle d'un promoteur obtenu par le procédé selon l'une quelconque des revendications 1 à 9.

20 13. Banque d'ADN dite banque de fusion, consistant en une banque d'ADN de *Kluyveromyces* sur laquelle a été réalisée une mutagénèse insertionnelle au moyen d'un transposon MiniMu comprenant un gène reporteur.

2693475

1/6

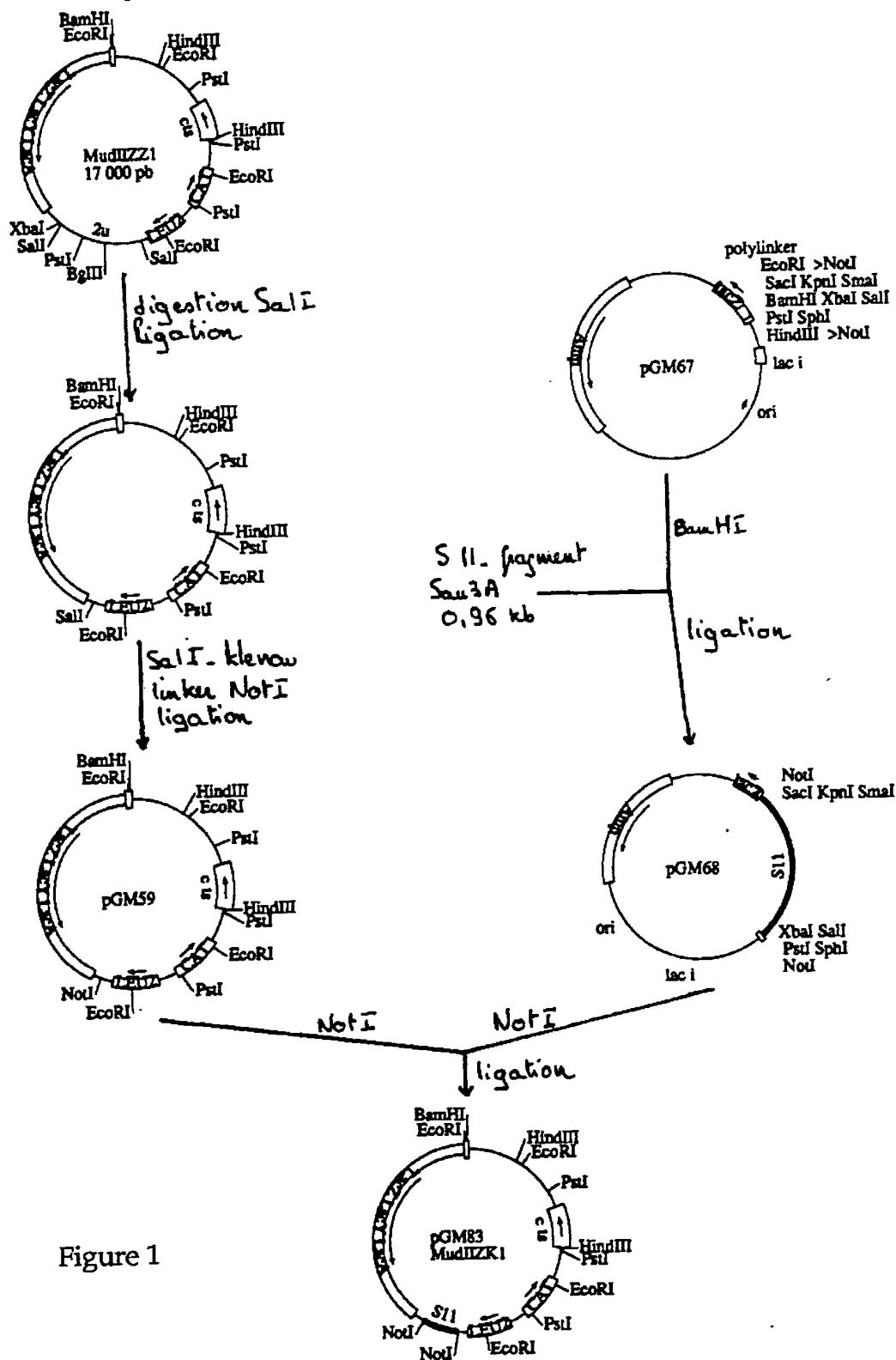


Figure 1

2693475

2/6

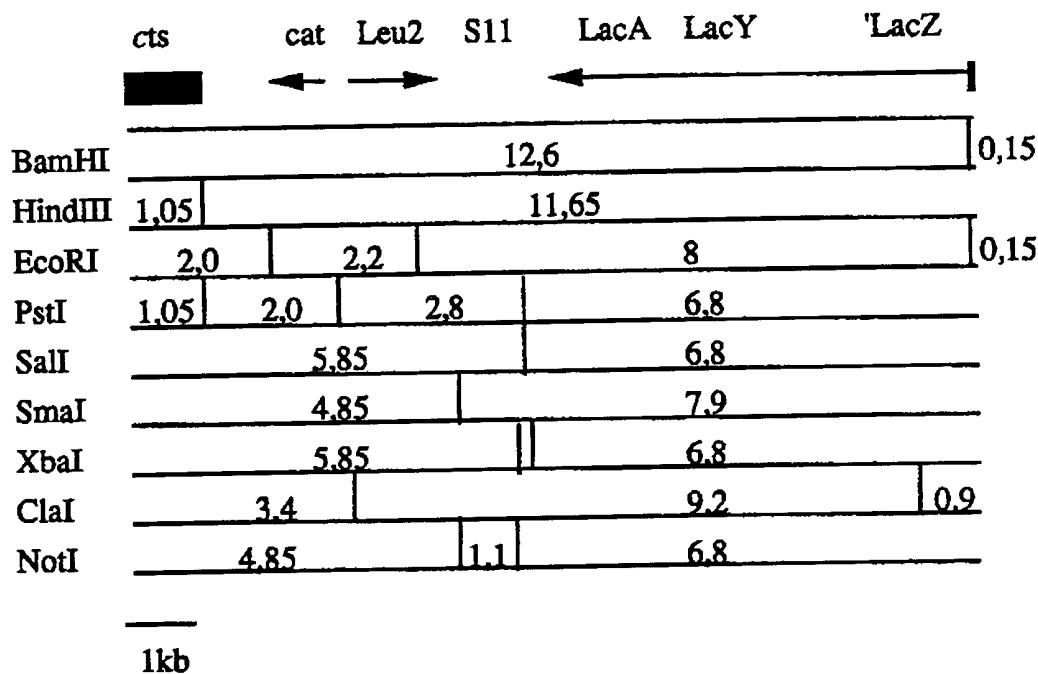


Figure 2

2693475

3/6

SEQ ID NO : 1

TYPE DE SEQUENCE : Nucléotide

LONGUEUR : 247 paires de bases

NOMBRE DE BRINS : 1

CONFIGURATION : Linéaire

TYPE DE MOLECULE : ADNc

ORIGINE : *Kluyveromyces lactis*

```
1 CATCGGCAAC TGGAGCCAAT GAGTATTGT CGTTAACCTTC ACCGTTGGT
51 CTACCCAAAT GAGAGGCCAA GACAATGGCC TTTGGCTTCT TTTCCAAAAC
101 GTATTGAATA GTAGGCAGAG CTGCTACAAT TCTTGTTG GAAGTGATCT
151 TCTTACCGTC CATGGCGTT GAAATCAACT CTGATGAACA CTCTTTACC
201 AGTGACATCC AAATCCTTGA CAGTCAATT TGAACTCAAA GACATTT
```

Figure 3

2693475**4/6****SEQ ID NO : 2**

TYPE DE SEQUENCE : Nucléotide

LONGUEUR : 212 paires de bases

NOMBRE DE BRINS : 1

CONFIGURATION : Linéaire

TYPE DE MOLECULE : ADNc

ORIGINE : *Kluyveromyces lactis*

1 CAGACAAGTC AGCATCGGGA CAGTTGGATT CATTGGACAA TTCACAGTAT
51 TCACAGGACA TACAAGAACCC GTTCAACCAT TTGATACCAG CAAAGTCACC
101 AATGTTCCAG CCCTTGACAT TTTCACCCAT AGCAACAACG ACACCAGCAC
151 CTTCGTGACC ACCAACTAAT GGCAATTGG TTGGCAAGGC AGTCACCCCTT
201 CCATGCGTGC AA

Figure 4

2693475

5/6

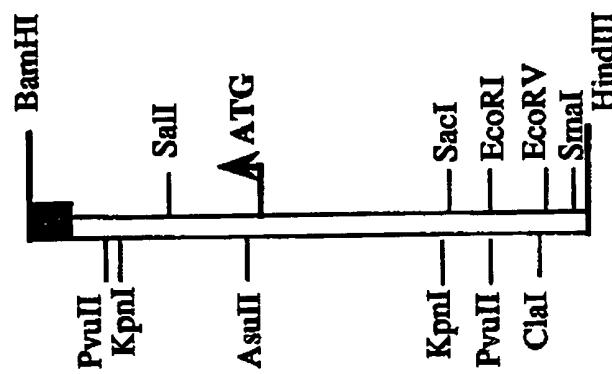


Figure 5

2693475

6/6

SEQ ID NO : 3

TYPE DE SEQUENCE : Nucléotide
LONGUEUR : 295aires de bases
NOMBRE DE BRINS : 1
CONFIGURATION : Linéaire
TYPE DE MOLECULE : ADNc
ORIGINE : *Kluyveromyces lactis*

1 ACAATAGAGC ACCGACCGAT AGAACCAAGT CAGCAGATTTC AACGGCTTCC
51 TTGACAGCTG GAGAAGATAG GGTACCGACG TAGACACCAC CGAATCTTGG
101 GTGCTTTTCG TCAATGGAAC CCTTACCCAT TGGGGTAACG AAGGCTGGGA
151 ATTGAGTCAA GTCGATCAAC TTCTTGGTCT CAGCCTTGGC ATCGTGTCTG
201 GAAACAACAAG CATCAGCCAA GATAACTGGG TTCTTAGCTT CCTTGATCAG
251 TTGCAAGACG TTTTCGATGA CTTCTCTTC GGCTTCTGGG TCATT

Figure 6

REPUBLIQUE FRANÇAISE

2693475

INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLE

RAPPORT DE RECHERCHE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FR 9208427
FA 473482

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
X	TRENDS IN BIOTECHNOLOGY vol. 6, no. 1, Janvier 1988, CAMBRIDGE GB pages 23 - 27 A. T. SCHAUER 'Visualizing gene expression with luciferase fusions' * Le document en entier * ---	1-9
A	WO-A-8 801 296 (THE GENERAL HOSPITAL CORPORATION) ---	-
A	EP-A-0 165 614 (HITACHI) ---	-
D,A	GENE. vol. 62, 1988, AMSTERDAM NL pages 45 - 54 B. DAIGNAN-FORNIER ET AL. 'In vivo functional characterization of a yeast nucleotide sequence: construction of a mini-Mu derivative adapted to yeast' ---	-
D,A	EP-A-0 361 991 (RHONE-POULENC SANTE) ---	DOMAINES TECHNIQUES RECHERCHES (Int. Cl.5)
A	EP-A-0 301 670 (GIST-BROCADES N.V.) -----	C12N C12Q
2		
Date d'achèvement de la recherche 07 AVRIL 1993		Emanation MOLINA GALAN E.
CATEGORIE DES DOCUMENTS CITES <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'un moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p>		
T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant		